

# Using syntactic and semantic features for classifying modal values in the Portuguese language

João Sequeira<sup>1</sup>, Teresa Gonçalves<sup>1</sup>, Paulo Quaresma<sup>1,4</sup>, Amália Mendes<sup>2</sup>, and Iris Hendrickx<sup>2,3</sup>

<sup>1</sup> Department of Informatics, University of Évora, Portugal

<sup>2</sup> Center for Linguistics of the University of Lisbon, Portugal

<sup>3</sup> Center for Language Studies, Radboud University Nijmegen, The Netherlands

<sup>4</sup> L2F – Spoken Language Systems Laboratory, INESC-ID, Portugal

**Abstract.** This paper presents a study made in a field poorly explored in the Portuguese language – modality and its automatic tagging. Our main goal was to find a set of attributes for the creation of automatic taggers with improved performance over the bag-of-words (bow) approach. The performance was measured using precision, recall and  $F_1$ . Because it is a relatively unexplored field, the study covers the creation of the corpus (composed by eleven verbs), the use of a parser to extract syntactic and semantic information from the sentences and a machine learning approach to identify modality values. Based on three different sets of attributes – from trigger itself and the trigger’s path (from the parse tree) and context – the system creates a tagger for each verb achieving (in almost every verb) an improvement in  $F_1$  when compared to the traditional bow approach.

## 1 Introduction

The automatic distinction between the factual and non-factual nature of events and the detection of the subjective perspective underlying texts is one of the concerns of the current trend in NLP that focuses on sentiment analysis and opinion mining. Modality is one such indicator of subjectivity and factuality in texts, as it is usually defined as the expression of the speaker’s opinion and of his attitude towards the proposition [16]. It traditionally covers epistemic modality, which is related to the degree of commitment of the speaker to the truth of the proposition (whether the event is perceived as possible, probable or certain), but also deontic modality (obligation or permission), capacity and volition.

The present experiments are related to the automatic tagging of modality for the Portuguese language, a topic that has received little attention for languages other than English. In fact, one of our goals is to be able to create a tagger using a small corpus sample to (semi) automatically tag a larger corpus with modality information. For this purpose, we use a corpus of 158.553 tokens, manually annotated with a modality scheme for Portuguese [8]. This paper restricts the

experiments to eleven modal verbs: *arriscar* (chance/risk/dare), *aspirar* (aspire), *conseguir* (manage to/succeed in/be able to), *considerar* (consider/regard), *dever* (shall/might), *esperar* (wait/expect), *necessitar* (need/require), *permitir* (allow/permit), *poder* (may/can), *precisar* (need) and *saber* (know). These verbs were selected, out of the total set of triggers of the annotated corpus, based on their polysemy: indeed, they all express two, or even three, modal meanings. This increases the difficulty of the automatic annotation process and makes them an excellent object of study for our experiments.

Our experiments in the automatic annotation of modality will first identify the modal verbs (which we call the modal trigger) and then assign them a modal value (out of the possible set of modal meanings that each verb expresses). A modal verb may be ambiguous between several readings. For instance, the verb *poder* may be

- **Epistemic**, stating that something is possible, as in example (1)
- **Deontic**, denoting a permission, as in (2), or
- may express an **Internal Capacity**, the fact that someone is able to do something, as in (3).

- (1) *O que é que vai fazer que julga poder ser marcante?*  
(What are [you] going to do that you believe may be significant?)
- (2) *Nenhum atleta devia poder estar nos Jogos Olímpicos depois de uma deserção e creio que nenhum país deve sentir-se satisfeito por exibir medalhas conquistadas por estrangeiros*.  
(No athlete should be allowed to be (lit: should can be) at the Olympic Games after defecting and I believe that no country should be pleased to exhibit medals won by foreigners.)
- (3) *Se a injustiça no mundo continuar, não sei até quando poderemos controlar as pessoas (...)*.  
(If world injustice goes on, [I] don't know till when we will be able to control people.)

To create the modality tagger, we first use a parser to get part-of-speech and syntactic information, then we identify modal triggers and, for those, we apply a machine learning approach to assign a modal value.

The paper is structured as follows: Section 2 presents related work on the automatic annotation of modality, Section 3 introduces the annotation scheme for Portuguese and presents statistics of the corpus used, Section 4 describes the features extracted and the system developed, and analyses the results obtained comparing them with a bag of words approach. Finally, in Section 5 we withdraw some conclusions and present some future work that would improve our system.

## 2 Related work

The annotation schemes for modality apply mostly to the English language, as in Baker *et al.* [2], Matsuyoshi *et al.* [9], Nirenburg and McShane [14] and

Sauriet *et al.* [19]. Two schemes have been devised for Portuguese: Hendrickx *et al.* [8] for written European Portuguese and Ávila *et al.* [21] for spoken Brazilian Portuguese. The existing schemes may focus on modality or integrate such information in a larger set of features that may cover, for instance, factuality, evidentiality, hedging, polarity and temporal information. These differences are discussed in detail in the work of Hendrickx *et al.* [8] and Nissim *et al.* [15]). Some of these schemes have been applied in experiments of automatic annotation and we focus here on reviewing the results that were obtained.

The experiments in Baker *et al.* [2] and Sauri *et al.* [19] report results with high success rates considering the complexity of the task at hand. Baker *et al.* [2] tested two rule-based modality taggers that identify both the modal trigger and its target and achieve results of 86% precision for a standard LDC data set. Sauri *et al.* [19] report, on the automatic identification of events in text and their characterization with modality features, to achieve accuracy values of 97.04% using the EviTA tool. Another experiment reported by Diab *et al.* [4] specifically addresses the annotation of belief by looking at auxiliary modal verbs. The authors do not consider the polysemy of modal verbs in their work and treat all auxiliary verbs as epistemic to avoid a source of noise in their system, although they acknowledge the fact that the verbs may have deontic meaning in certain contexts (note that our experiment, reported in Section 4, deals with the added complexity of multiple modal meanings). Prabhakaran *et al.* [17] extend the experiment on belief with tests on tagging different modality values (**Ability**, **Effort**, **Intention**, **Success** and **Want**); the authors report experiments on two very different annotated corpora: MTurk data composed of email threads and using only those examples for which at least two Turkers agreed on the modality and the target of the modality; and on a gold dataset, that contains sentences from newswire, letters and blogs in addition to emails. The results differ greatly according to the corpus used: the MTurk data achieves an overall 79.1% F-measure while the gold dataset presents 41.9% F-measure. Furthermore, there was a specific shared task at CoNLL2010 [5] on the detection of uncertainty and its linguistic scope by identifying hedging cues, which includes a broader set of lexical and syntactic clues when compared to modality as discussed in this paper. Finally, the annotation of events in the area of BioNLP includes in some cases values related to modality and factuality. The system described in [12] seeks to label events with the dimension 'level of certainty' and attains F-measures of 74.9% for 'low confidence' and 66.5% for 'high but not complete confidence'.

While our experiment focuses on auxiliary modal verbs but also main verbs with modal meaning, the work of Ruppenhofer and Rehbein [18] consider only the five auxiliary English verbs can/could, may/might, must, ought, shall/should. The authors predict the verb's modal value in context by training a maximum entropy classifier on features extracted from the corpus and improve the baseline for all verbs (but must), achieving accuracy numbers between 68.7% and 93.5%.

### 3 Modality Corpus

Our experiment applies over a corpus annotated with the annotation scheme for Portuguese presented in [8]. This scheme takes the concept of Modality as the expression of the speaker’s attitude towards the proposition, so the concept of factuality is not included, contrary to approaches such as [15], who accounts for both values but in different layers of the annotation scheme. Tense and mood are also categories that are not taken into account, despite their relation with modality. The authors report that the approach is similar to the OntoSem ([10]) annotation scheme for modality [14]. Finally, the annotation is not restricted to modal verbs and covers also nouns, adjectives and adverbs, although for this experiment we only focus on a specific set of verbs.

Next subsections introduce the annotation scheme and corpus used for this work.

#### 3.1 Annotation scheme

Several modal values are included based not only on the modality literature but also on studies focused on annotation and information extraction (e.g. [16,1,2]). Seven main modal values are considered: **Deontic**, **Effort**, **Epistemic**, **Evaluation**, **Participant-internal**, **Success** and **Volition**. Some of these values are further classified into sub-values:

- Deontic modality has two subvalues: **Obligation** and **Permission**. This includes what is sometimes considered Participant-external modality, as in [1];
- Epistemic modality is further divided in **Knowledge**, **Belief**, **Doubt**, **Possibility** and **Interrogative**. Contexts traditionally considered of the modal type ‘evidentials’ (i.e., supported by evidence) are annotated as **Epistemic Belief**;
- Participant-internal modality has two sub-values: **Necessity** and **Capacity**.

The annotation scheme comprises several components: the trigger, which is the lexical element conveying the modal value; its target; the source of the event mention (speaker or writer); and the source of the modality (agent or experiencer). The trigger receives an attribute modal value, while both trigger and target are marked for polarity. For example, the modal verb *poder* in sentence (4) is underlined.

- (4) *Caso a avaliação seja positiva, a empresa de recauchutagem poderá salvar grande parte do equipamento que se encontra no interior das instalações, garantindo assim a laboração num curto espaço de tempo, que, segundo o administrador da empresa, António Santos, não poderá exceder os 15 dias.* (If the evaluation is positive, the retreading production unit may save most of the equipment in the company premises, and so guaranty that the operation activity is done in a short span of time, that, according to the administrator of the company, António Santos, may not exceed 15 days.)

This sentence contains two other triggers: a first occurrence of the same verb *poderá* expresses the modal value Epistemic possibility, the trigger *garantido* (to guaranty) expresses Epistemic belief. We focus on the annotation of the second trigger *poder* in more detail. The target is discontinuous and we mark it here with the symbol @, although it is expressed in XML in our editor.

- Trigger: *poderá*
- Modal value: Deontic\_permission. Polarity: negative
- Target: *a laboração@exceder os 15 dias*
- Source of the modality: *António Santos*
- Source of the event: writer
- Ambiguity: none

Full details on the annotation scheme and on the results of an inter-annotator experiment are provided in [8]. An enriched version with the interaction between Focus and Modality, specifically the case of exclusive adverbs, is presented in [11].

### 3.2 Corpus

The annotation scheme was applied to a corpus sample extracted from the written subpart of the Reference Corpus of Contemporary Portuguese (CRPC) [6]. Details about the selection of the sample are provided in [8]. The MMAX2<sup>5</sup> annotation software tool [13] was used for the manual annotation task. The elements of the annotation consist of markables that are linked to the same modal event, which is called a set.

For this study we used a subset of the annotated corpus by including the sentences from eleven verbs. Table 2 resumes the information about each verb.

**Table 1.** Corpus characterization: number of sentences per modal value for each verb.

verb	modal values	number of sentences			
		total	val_1	val_2	val_3
arriscar	2	44	19	25	
aspirar	2	50	31	19	
conseguir	2	84	41	43	
considerar	2	29	18	11	
dever	2	108	37	71	
esperar	2	52	26	26	
necessitar	2	50	8	42	
permitir	2	78	60	18	
poder	3	236	42	154	40
precisar	2	54	45	9	
saber	2	103	93	10	

<sup>5</sup> The MMAX2 software is platform-independent, written in java and can freely be downloaded from <http://mmax2.sourceforge.net/>.

## 4 Developed system

The developed system works in two steps: first, it identifies the modal verbs and then it labels the appropriate modal value in its specific context. Modal verbs are identified by the automatic analysis of the output of the syntactic parser and the modal values are labelled using a Machine Learning approach.

The syntactic analysis is performed using the PALAVRAS parser [3] and the set of sentences that include modal verbs are selected to build the data for the Machine Learning algorithm.

### 4.1 Experimental setup

Most machine learning algorithms use the vector space model to represent the input data. Using this approach each sentence needs to be transformed into a set of features that can be boolean, nominal or real valued. This work uses the output of PALAVRAS to build those features: we include information from the trigger itself, from the syntactic tree path and from the trigger's context. Besides this approach a bag-of-words representation of the sentences was also considered as a baseline representation. Next subsection describes in detail the information extracted from the syntactic tree and how it was represented.

The SVM (Support Vector Machine) algorithm [20] was chosen to label the modal value of each verb. Several initial experiments were conducted with different degrees of the polynomial kernel ( $n \in 1, 2, 3$ ) and values of the C parameter ( $C \in 0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000, 10000$ ). Those experiments enabled us to chose the linear kernel with  $C = 1$ .

Different sets of extracted attributes were evaluated and compared with a typical bag-of-words approach. For the evaluation we used a 5-fold stratified cross-validation procedure (repeated twice) and computed average precision, recall and F1 performance measures. Appropriate statistical tests with 95% of significance were applied to analyse the differences between results.

These machine learning experiments were conducted using Weka framework [7].

### 4.2 Feature extraction

The information extracted to build the attributes is inspired in the work by Ruppenhofer and Rehbein [18]. Their approach used three specific sets of attributes: information from the **trigger**, from the **path** (from the trigger to the root taken from the syntatic parse tree) and from the **context** around the trigger.

Using the PALAVRAS parser we also extracted attributes from the trigger, context and path including all possible information given by the parser output:

- for the trigger: besides the trigger itself we included info from the ancestral nodes (father, grandfather and great grandfather);
- for the path: besides collecting info from the trigger to the root, we also included info from the its left and right nodes;

- for the context: we collected info about the previous and following words using different size windows.

Table 2 summarizes the attributes extracted and next sub-sections detail the extracted attributes and summarize them.

**Table 2.** Attributes extracted from trigger, path and context.

<b>trigger</b>		<b>path</b>		<b>context</b>	
source	attributes	source	attributes	source	attributes
trigger	POS	siblings	POS	left/right	POS
	function		function	trigger	word
	role		role		lemma
	morphological semantic		morphological semantic		
ancestors	POS	trigger to root	POS		
	function		function		

**Trigger related attributes.** For each trigger word we extracted the POS tag, function, morphological and semantic information, and the role (if it exists). For the ancestral nodes (father, grandfather and great-grandfather) we extracted the POS tag and function. All information was represented as binary attributes (present/not present).

**Path related attributes.** For each trigger we extracted POS tags and functions from the tree’s path (all nodes from the trigger word to the root) and also the POS tag, function, morphological and semantic information, and the role (if exists) from the path of sibling nodes (left and right). All information was represented as numerical attributes (counts over each possible value).

**Context related attributes.** For each trigger we extracted information about the POS tags, words and lemmas in the surrounding context with a size window equal to five words (with the trigger word in the middle). All information was represented as numerical attributes (counts over each possible value).

**Datasets characterization.** As already mentioned, besides building a representation using the output of PALAVRAS, a traditional bag-of-words approach was also considered. Table 3 resumes, for each verb, the number extracted for each specific set of attributes.

**Table 3.** Datasets characterization: number attributes for bag-of-words, trigger, path and context sets.

verb	bow	trigger	path	context
arriscar	642	577	734	2548
aspirar	704	600	775	2736
conseguir	1093	649	838	4170
considerar	552	534	678	1918
dever	1578	645	845	5596
esperar	611	554	703	2358
necessitar	714	576	747	2732
permitir	1320	685	900	4966
poder	2736	715	940	9770
precisar	588	574	728	2320
saber	1237	696	912	4504

### 4.3 Experiments

In order to evaluate the discrimination power of each set of attributes, eight experiments were done: bag-of-words, trigger, path, context, trigger+path, trigger+context, path+context and trigger+path+context. Tables 4, 5 and 6 present the precision, recall and F1 values, respectively. Values statistically different (better or worse) from the corresponding bag-of-words experiment are boldfaced.

**Table 4.** Results: precision values

	bow	trigger	path	context	tg+pth	tg+ct	pth+ct	all
arriscar	.638	.686	.771	.757	.719	.750	.833	.804
aspirar	.741	.853	.795	.694	.778	.756	.778	.779
conseguir	.540	.583	.595	.678	.592	.672	.714	.684
considerar	.402	.489	.526	.611	.582	.536	.650	.660
dever	.700	.662	.568	.626	.636	.692	.611	.602
esperar	.745	.610	<b>.527</b>	.595	.545	.619	.477	.577
necessitar	.708	.723	.735	.732	.709	.701	.690	.698
permitir	.593	.666	<b>.754</b>	<b>.785</b>	.702	<b>.786</b>	<b>.812</b>	<b>.811</b>
poder	.530	.486	.529	.522	.544	.520	.484	.536
precisar	.698	.700	.669	.757	.700	.788	.736	.736
saber	.815	<b>.906</b>	.833	.861	.881	.903	.843	<b>.917</b>

### 4.4 Discussion of results

For the bag-of-words approach we got results ranging between 0.402 (*considerar*) and 0.815 (*saber*) for precision; 0.530 (*conseguir*) and 0.903 (*saber*) for recall; and 0.472 (*conseguir*) and 0.857 (*saber*) for F1.



**Table 5.** Results: recall values

	bow	trigger	path	context	tg+pth	tg+ct	pth+ct	all
arriscar	.614	.672	.726	.740	.669	.718	.800	.765
aspirar	.710	.820	.76	.710	.770	.730	.770	.760
conseguir	.530	.565	.59	.671	.577	.647	.696	.660
considerar	.587	.510	.513	.647	.547	.513	.657	.623
dever	.700	.652	.601	.645	.611	.690	.626	.607
esperar	.720	.605	.527	.595	.547	.606	.499	.577
necessitar	.840	.780	.73	.830	.740	.790	.720	.770
permitir	.770	.704	.744	.795	.698	.795	.827	.807
poder	.650	.574	.61	.581	<b>.574</b>	.568	<b>.551</b>	.579
precisar	.835	.739	.732	.826	.750	.827	.799	.817
saber	.903	.898	.869	.913	.903	.937	.908	.942

Looking at Table 2 we are able to find that *considerar*, *conseguir* and *saber* have 29, 84 and 103 examples, respectively (with 2 possible modal values); while *considerar* and *conseguir* constitute balanced datasets, for *saber* there is a ratio of 9:1. From this we can say that *saber* seems to be a easier verb to assign the modal value and that the unbalanced dataset does not seem hurt the performance.

**Table 6.** Results: F1 values

	bow	trigger	path	context	tg+pth	tg+ct	pth+ct	all
arriscar	.605	.660	.712	.736	.658	.708	.794	.758
aspirar	.693	.823	.759	.694	.770	.730	.763	.757
conseguir	.552	.554	.581	.670	.563	.640	.689	.650
considerar	.472	.470	.510	.611	.543	.489	.634	.618
dever	.668	.644	.566	.617	.609	.685	.610	.597
esperar	.712	.599	.513	.568	.537	.596	.473	.562
necessitar	.768	.746	.724	.775	.718	.742	.699	.728
permitir	.670	.672	.742	.782	.683	.775	<b>.807</b>	.794
poder	.540	.520	.535	.537	.548	.538	.510	.550
precisar	.760	.716	.697	.787	.721	.799	.763	.770
saber	.857	.895	.849	.884	.889	.914	.872	<b>.923</b>

From Table 4 its possible to state that, for most experiments, there are no statistical differences with the bag-of-words approach. Nevertheless, when using information extracted from PALAVRAS, we got improvements in precision values over the verbs *permitir* and *saber* when compared with the bag-of-words approach. For *permitir* we got better results when using path (0.754), context (0.785), trigger+context (0.786), path+context (0.812) and all (0.811) features; for *saber* better results were obtained with trigger (0.906) and all (0.917) attributes. If we look at the number of features of each setting (Table 3) we can

conclude that for these verbs we are able to get higher precision using less attributes: for *permitir*, we have 747 vs. 1320 attributes for path vs. bag-of-words; for *saber*, we have 696 vs. 1237 attributes for trigger vs. bag-of-words. On the other hand, the verb *esperar* presents worse results when compared to bag-of-words for path (0.527) experiments.

For recall values (Table 5) we got even less statistical differences. Only verb *poder* got different results and they were worse (using trigger+path attributes with a recall of 0.574 and for path+context ones with a recall of 0.551).

Finally, and looking at Table 6, we can state that using information from the parse tree, we were able to improve the F1 values for 2 verbs: for *permitir* with the path+context setting (0.807) and for *saber* with all attributes (0.923).

## 5 Conclusions and future work

With this work, we tried to address a topic that has not been much explored in the Portuguese language – the automatic tagging of modality. The correct tagging of modality is important since it is linked to the current trend in NLP on sentiment analysis and opinion mining.

Due to this limited research it was necessary to implement a viable corpus for the work to be done: 11 verbs were chosen to be studied and the parser PALAVRAS was used to obtain morphological, syntactic and some semantic information from the sentences. Following a similar approach to Ruppenhofer and Rehbein [18] we defined three sets of attributes: trigger related, path related and context related ones.

Using Weka framework we conducted several experiments to study the effect on the usage of linguistic information to identify modal values. As baseline we used a bag-of-words approach and calculated precision, recall and  $F_1$  measures. While we were able to get better results (precision and F1) with some settings for some verbs (*permitir* and *saber*), most experiments, even with higher performance values, are not significantly different from the bag-of-words approach (mainly because of the small number of training examples). Also, we obtained worse precision results for verb *esperar* using path attributes and worse recall values for verb *poder* using trigger+path and path+context ones.

With this work we were able to find a set of linguistic information related attributes that can be used to identify the modal values for the Portuguese language and from the results we can conclude that the use of information extracted from the parse trees does not harm the performance of the automatic taggers and can even, for some verbs and combinations enhance it.

Considering that our training corpus was relatively small and that we selected challenging verbs in our experiment, we believe that our goal, of creating a larger corpus with modal information by a (semi) automatic tagging process, could lead to positive results in the future. We plan to study the role played by each feature in our system and to observe in more detail the reasons why some verbs reach higher scores than others.

As we are currently applying a 'word expert' approach and training separate classifiers for different verbal triggers, it is clear that this approach will not be able to handle modal triggers that were not seen before. We intend to study this problem and try to train, for example, a general modal trigger classifier that is not dependent on the verb itself. Also, besides developing the tagger for the trigger we intend to build a system that is able to identify the target of the modality.

## Acknowledgements

This work was partially supported by national funds through FCT – Fundação para a Ciência e Tecnologia, under project Pest-OE/EEI/LA0021/2013 and project PEst-OE/LIN/UI0214/2013.

## References

1. der Auwera, J.V., Plungian, V.A.: Modality's semantic map. *Linguistic Typology* 1(2), 79–124 (1998)
2. Baker, K., Bloodgood, M., Dorr, B., Filardo, N.W., Levin, L., Piatko, C.: A modality lexicon and its use in automatic tagging. In: Chair), N.C.C., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., Tapias, D. (eds.) *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. European Language Resources Association (ELRA), Valletta, Malta (may 2010)
3. Bick, E.: *The parsing system PALAVRAS*. Aarhus University Press (1999)
4. Diab, M.T., Levin, L.S., Mitamura, T., Rambow, O., Prabhakaran, V., Guo, W.: Committed belief annotation and tagging. In: *Third Linguistic Annotation Workshop*. pp. 68–73. The Association for Computer Linguistics, Singapore (August 2009)
5. Farkas, R., Vincze, V., Móra, G., Csirik, J., Szarvas, G.: The conll-2010 shared task: Learning to detect hedges and their scope in natural language text. In: *Proceedings of the Fourteenth Conference on Computational Natural Language Learning*. pp. 1–12. Association for Computational Linguistics, Uppsala, Sweden (July 2010)
6. Génèreux, M., Hendrickx, I., , Mendes, A.: Introducing the reference corpus of contemporary portuguese on-line. In: Calzolari, N., Choukri, K., Declerck, T., Dogan, M.U., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S. (eds.) *LREC'2012 – Eighth International Conference on Language Resources and Evaluation*. pp. 2237–2244. European Language Resources Association (ELRA), Istanbul, Turkey (May 2012)
7. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The weka data mining software: An update. *SIGKDD Explor. Newsl.* 11(1), 10–18 (Nov 2009)
8. Hendrickx, I., Mendes, A., Mencarelli, S.: Modality in text: a proposal for corpus annotation. In: Chair), N.C.C., Choukri, K., Declerck, T., Dogan, M.U., Maegaard, B., Mariani, J., Moreno, A., Odijk, J., Piperidis, S. (eds.) *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*. European Language Resources Association (ELRA), Istanbul, Turkey (may 2012)

9. Matsuyoshi, S., Eguchi, M., Sao, C., Murakami, K., Inui, K., Matsumoto, Y.: Annotating event mentions in text with modality, focus, and source information. In: Chair), N.C.C., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., Tapias, D. (eds.) *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. European Language Resources Association (ELRA), Valletta, Malta (may 2010)
10. McShane, M., Nirenburg, S., Beale, S., O'Hara, T.: Semantically rich human-aided machine annotation. In: *Proceedings of the Workshop on Frontiers in Corpus Annotations II: Pie in the Sky*. pp. 68–75. Association for Computational Linguistics, Ann Arbor, Michigan (June 2005)
11. Mendes, A., Hendrickx, I., Salgueiro, A., Ávila, L.: Annotating the interaction between focus and modality: the case of exclusive particles. In: *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*. pp. 228–237. Association for Computational Linguistics, Sofia, Bulgaria (August 2013)
12. Miwa, M., Thompson, P., McNaught, J., Kell, D.B., Ananiadou, S.: Extracting semantically enriched events from biomedical literature. *BMC Bioinformatics* 13, 108 (2012)
13. Müller, C., Strube, M.: Multi-level annotation of linguistic data with MMAX2. In: Braun, S., Kohn, K., Mukherjee, J. (eds.) *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods*, pp. 197–214. Peter Lang, Frankfurt a.M., Germany (2006)
14. Nirenburg, S., McShane, M.: Annotating modality. Tech. rep., University of Maryland, Baltimore County, USA (March 2008)
15. Nissim, M., Pietrandrea, P., Sanso, A., Mauri, C.: Cross-linguistic annotation of modality: a data-driven hierarchical model. In: *Proceedings of IWCS 2013 WAMM Workshop on the Annotation of Modal Meaning in Natural Language*. pp. 7–14. Association for Computational Linguistics, Postam, Germany (2013)
16. Palmer, F.R.: *Mood and Modality*. Cambridge textbooks in linguistics, Cambridge University Press (1986)
17. Prabhakaran, V., Bloodgood, M., Diab, M., Dorr, B., Levin, L., Piatko, C.D., Rambow, O., Van Durme, B.: Statistical modality tagging from rule-based annotations and crowdsourcing. In: *Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics*. pp. 57–64. ExProM '12, Association for Computational Linguistics, Stroudsburg, PA, USA (2012)
18. Ruppenhofer, J., Rehbein, I.: Yes we can!? annotating english modal verbs. In: Chair), N.C.C., Choukri, K., Declerck, T., Doğan, M.U., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S. (eds.) *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*. European Language Resources Association (ELRA), Istanbul, Turkey (may 2012)
19. Sauri, R., Verhagen, M., Pustejovsky, J.: Annotating and recognizing event modality in text. In: *FLAIRS Conference*. pp. 333–339 (2006)
20. Vapnik, V.N.: *Statistical Learning Theory*. Wiley-Interscience (1998)
21. Ávila, L., Melo, H.: Challenges in modality annotation in a brazilian portuguese spontaneous speech corpus. In: *Proceedings of IWCS 2013 WAMM Workshop on the Annotation of Modal Meaning in Natural Language*. Association for Computational Linguistics, Postam, Germany (2013)